ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Computer Science
Procedia Computer Science 12 (2012) 96 – 101

Complex Adaptive Systems, Publication 2
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2012- Washington D.C.

Hybrid Sampling Strategy-based Multiobjective Evolutionary Algorithm

Wenqiang Zhang[a], Lin Lin[bc], Mitsuo Gen[cd]* and Chen-Fu Chien[d]

[a]*College of Information Science and Engineering, Henan University of Technology, China*
[b]*School of Software, Dalian University of Technology, China*
[c]*Fuzzy Logic Systems Institute, Japan*
[d]*Dept. of Industrial Engineering & Engineering Management, National Tsing Hua University, Taiwan*

Abstract

Recently more research works are focused on multiobjective evolutionary algorithm (MOEA) duo to its ability of global and local search for solving multiobjective optimization problem (MOOP) and ability to provide more practical solutions to decision maker; however, most of existing MOEAs cannot achieve satisfactory results in both quality and computational speed. This paper proposes a hybrid sampling strategy-based multiobjective evolutionary algorithm (HSS-MOEA) to deal with such problem. HSS-MOEA tactfully combines the sampling strategy of vector evaluated genetic algorithm (VEGA) and the sampling strategy according to a new Pareto dominating and dominated relationship-based fitness function (PDDR-FF). The sampling strategy of VEGA prefers the edge area of the Pareto front and PDDR-FF-based sampling strategy has the tendency converging toward the central area of the Pareto front. The hybrid sampling strategies preserve both the convergence rate and the distribution performance. Numerical comparisons show that HSS-MOEA could get the better convergence performance, slightly better or equivalent distribution performance, and obviously better efficiency than existing MOEAs.

Keywords: evolutionary algorithm; hybrid sampling; multiobjective optimization

1. Introduction

MOOP is a practical, important but very intractable optimization problem in which two of more conflicting objectives should be considered together, and many Pareto-optimal solutions with incommensurable quality are generated for decision makers. Decision makers like the multiple solutions with same good quality so that they could select ones according to their subjective preferences. In the most single objective optimization problem, more problem-dependent considerations are concerned and one global optimal solution and less computation time are desired. However MOOP needs additional cares in convergence mechanism of finding sufficient number of global Pareto optimal solutions as soon as possible, and distribution mechanism of distributing them as evenly as possible. Moreover, reduction of computation time becomes more difficult accordingly.

MOEAs have been recognized to be well-suited for solving MOOPs [1-3]. VEGA divides the population into m

* Corresponding author. Tel.: +82-90-9793-1826
E-mail address: gen@flsi.or.jp

 doi:10.1016/j.procs.2012.09.037

sub-populations (*m* is number of objectives), each of which evolves toward a single objective [4]. Weight sum methods transform MOOP into single-objective problem by assigning different weight for each objective and aggregating them. However, fixed weight or random weight could cause search bias in evolving process. Gen and Cheng propose an adaptive weights genetic algorithm (AWGA) which utilizes some useful information from current population to readjust weights in order to obtain a search pressure towards to positive ideal point [5]. Lin *et al* propose an interactive adaptive weight GA (IAWGA) by using AWGA and Pareto ranking based approach to classify the difference among non-dominated solutions (or dominated solutions) clearly [3].

As two classical MOEAs, NSGA-II [6] and SPEA2 [7] have been proven to get better quality in solving MOOPs. Zhang and Fujimura proposed an improved vector evaluated genetic algorithm with archive (IVEGA-A) that combined VEGA and Pareto-based scale-independent fitness function (GPSI-FF)-based archive mechanism [10].

To increase the quality (both convergence and distribution) and to reduce the computational time, a new HSS-MOEA is proposed. A novel PDDR-FF is proposed to evaluate the individuals. The proposed method deliberately combines the two different mechanisms. One is the sampling strategy of VEGA with a preference for the edge region of the Pareto front, and the other is the sampling strategy of PDDR-FF with tendency converging toward the central area of the Pareto front. These two mechanisms not only preserve the convergence rate, but also guarantee the better distribution performance.

The paper is organized as follows: Section 2 reviews the literature of recent research works; Section 3 presents the detailed HSS-MOEA approach; Section 4 gives a discussion and analysis of numerical experiments results; finally, the conclusion and future work are given in Section 5.

2. Literature Review

As the first one of MOEAs, VEGA just selects individuals into mating pool for one objective in selection phase. The benefit of VEGA is the strong ability to converge to the edge region of the Pareto front by its simple sampling strategy and less time complexity. However, the qualities (especially, diversity) of VEGA are not good because of the selection bias.

In AWGA, positive ideal point and negative ideal point are defined according to the current population so as to the approximate objective function value interval can be calculated. The weight for each objective can be defined by ideal point that it can be changed with the evolving process. AWGA uses these adaptive weights to normalize objective values and accumulate them as fitness value. As a result, AWGA has tendency towards center region of Pareto front because the individuals locating at center area have bigger fitness value than edge area.

Lin and Gen propose an IAWGA which combines the AWGA based fitness value and Pareto ranking strategy. First AWGA fitness value is calculated normally, then dominating and dominated relationship of each individuals are calculated accordingly. For minimization problem, AWGA fitness value is added 1 if the individual belongs to nondominated solution; otherwise, only add 0 to the AWGA fitness value. Obviously, IAWGA has same disadvantage as AWGA that individuals locating at center area of Pareto front are bigger than edge area' and all of nondominated individuals are allocating along Pareto front with rank 1.

In NSGA-II, Pareto ranking need to be calculated, then sorts these individuals according to ranking value by ascending. In Pareto ranking phase, first to calculate the dominance relationship of each individual, those individuals without dominated by anyone are addressed as rank 1, then deletes those individuals with rank 1 and recalculates the dominance relationship of left individuals to get rank 2, and so on. After Pareto ranking phase, the crowding distance need to be calculated for each individual. When updating the archive, the individuals with rank 1 should be inserted first, then rank 2. If rank *r* cannot be fully inserted into archive, insert the individuals in the descending order of the crowding distance until the archive is full. However, calculation of rank and crowding distance need much CPU time.

In SPEA2, first to calculate the individual's strength, which is the number of individuals it dominates in population and archive. Then to get the individual's raw fitness, which is the total strengths value of individuals dominate it. Also need to calculate the individual's distance to each individual. When updating the archive, both of the raw fitness and the distance are considered. Duo to the complicated computation of distance value and pruning scheme when updating archive cause that SPEA2 need much CPU time than NSGA-II but better distribution performance than NSGA-II.

Ho *et al*. propose a generalized Pareto-based scale-independent fitness function (GPSI-FF) to solve the large parameter optimization problem [8]. The GPSI-FF can obviously speed up the convergence rate [9], especially around the central area of the Pareto front. The aforementioned VEGA has preference for the edge region of Pareto front. Therefore it is reasonable to hybridize these two methods to achieve better convergence performance approaching to central and edge areas of whole Pareto font. According to these ideas, Zhang and Fujimura propose IVEGA-A in which both VEGA and GPSI-FF-based archive not only preserve the convergence rate but also guarantee better distribution performance. However, the difference between nondominated and dominated individuals can be decreased by GPSI-FF values. This disadvantage causes that more dominated individuals could be held in archive (external population) while nondominated ones would be removed from archive. It reduces the performance of archive, unless an enough large size of its archive is set to store a sufficient number of individuals.

3. Hybrid Sampling Strategy-based Multiobjective Evolutionary Algorithm

3.1. Pareto dominating and dominated relationship-based fitness function

For covering the disadvantage of GPSI-FF, a new PDDR-FF-based fitness function is proposed to evaluate the individuals. The PDDR-FF of an individual s_i is calculated by the following function:

$$eval(s_i)=q(s_i)+1/(p(s_i)+1),\ i=1,2,\ldots,popSize \qquad (1)$$

where $p(s_i)$ is the number of individuals, that can be dominated by the individual s_i, $q(s_i)$ that can dominate the individual s_i. The smaller value is better. According to the Eq. (1), PDDR-FF can set the obvious difference values between the nondominated and dominated individuals. If the individual belongs to nondominated one, its fitness value will not exceed one. The fitness value of individual which is dominated by other will exceed one. Furthermore, ones with different numbers of dominating are also given the different fitness values even though they are all nondominated individuals. It is obvious that the nondominated individuals locating around the central region of Pareto font with bigger domination area will have smaller values (near to 0) than the edge points (near to 1). Therefore, PDDR-FF gives the sensible difference values between the nondominated and dominated individuals. Moreover, ones with different numbers of dominating are also set as different fitness values even though they are all nondominated individuals. The individuals locating around the central region of Pareto font will have smaller values than the edge points.

Duo to only calculation dominating and dominated relationship among individuals without calculating distance makes PDDR-FF much faster than other approaches with diversity preservation mechanism.

3.2. Main Framework of HSS-MOEA

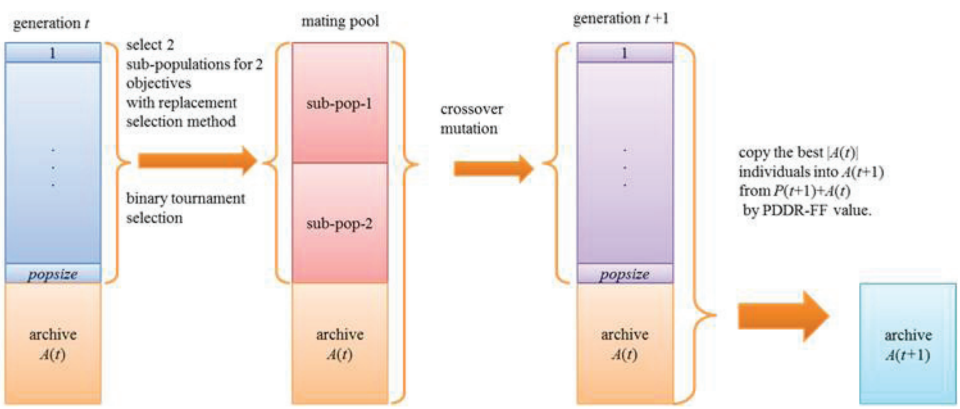

Fig. 1: The evolving process in one generation

The evolving process of one generation of HSS-MOEA is shown in Fig. 1. $A(t)$ represents the archive as an external population at generation t and $P(t)$ represents the population at generation t. In selection phase of HSS-MOEA, the PDDR-FF based sampling strategy has the advantage with the tendency converging toward the centre area of the Pareto front, but drawback to the edge region. It causes bad distribution performance. The sampling strategy of VEGA prefers the edge rather than certer regions of Pareto front that it causes VEGA cannot achieve better distribution performance. So it is natural, reasonable and possible to combine these two methods to improve the overall performance and reduce the computation time of the algorithm.

The solution procedure of one generation includes 3 phases.

Phase 1: Generation of Mating Pool (Hybrid Sampling)

Step 1: Selection 1 (Sampling Strategy of VEGA)

In this step, the sampling strategy of VEGA is used to select the better individuals into part of mating pool. For two objectives optimization problem, individuals are selected with replacement according to objective 1 into sub population 1 while ignoring objective 2 until the size of the sub population 1 (half of population size) is reached. In the same manner, individuals are selected for objective 2 into sub population 2 without considering objective 1 until sub population 2 is full. As a result, the individuals locating at edge areas of Pareto front have higher probability to be selected into sub populations.

Step 2: Selection 2 (PDDR-FF based Elitist Sampling Strategy)

In this study, the sub populations and *A(t)* are combined to form a mating pool. In the mating pool, sub-pop-1 stores the good individuals for one objective, and sub-population 2 holds the good individuals for the other objective. These two sub populations ensure that the solutions approach to edge areas of true Pareto front. The archive saves the individuals with good PDDR-FF values that guarantee the asolutions converge to center area of true Pareto front. As two objectives optimization problem, the sizes of those two sub populations and archive are set as the same as half the population size. Therefore, in the mating pool, one-third of the individuals serve one objective, one-third the other objective, and the left one-third both the two objectives. The archive mechanism tries to cover the selection bias of VEGA. These three parts of the mating pool make the solutions converge to the true Pareto front evenly.

Phase 2: Reproduction

Some crossover and mutation operators are used to reproduce new chromosomes.

Phase 3: Archive Maintenance

The individuals of *A(t)* and *P(t)* are combined to form a temporary archive *A'(t)*. Thereafter, the PDDR-FF values of all individuals in *A'(t)* are calculated and sorted in a ascending order. The smallest $|A(t)|$ individuals in *A'(t)* are copied to form *A(t+1)*. This archive updating mechanism likes a elitist sampling strategy to keep the better individuals with better PDDR-FF values.

The strong convergence capability of VEGA and PDDR-FF ensures that the HSS-MOEA has the ability to converge to the true Pareto front both in central and edge regions. The preferences for the edge area of the Pareto front in VEGA and the central area of the Pareto front in PDDR-FF guarantee that the HSS-MOEA distributes along the Pareto front evenly. Moreover, HSS-MOEA has less computing time.

4. Experiments and Discussion

Although there are many benchmark problems which have been usually used to check the efficacy and efficiency of MOEA, a most complicated but practical scheduling problem, processping planning and scheduling (PPS) is adopted in this study. PPS is to process a set of prismatic parts into completed products effectively and economically in a manufacturing system. Process planning (PP) is the determination of optimal process plans, i.e. operations (machine, tool, tool access direction) and their sequences. The scheduling is determination of the most appropriate moment to execute each operation with competitive resources. The minimizing makespan and minimizing variation of workload for each machine are used as the two objectives in PPS problems and 4 parts (each part has 20, 16, 14 and 7 operations, respectively) problem are used as simulation experimental data(see [9][10] for details).

All the simulation are performed on AMD sempron X2 180 processor (2.40 GHz clock) and 2GB memory. The adopted parameters are listed as follows: population size, 100; maximum generation, 500; archive size, 50; crossover probability, 0.70; mutation probability, 0.30. HSS-MOEA, IVEGA-A, VEGA, AWGA, IAWGA, NSGA-II, and SPEA2 are run 30 times to compared the results with each other. It should be noted that the parameters of all 7 methods are the same, except for the size of archive. The archive sizes of HSS-MOEA and IVEGA-A are set to be half the population size, 50, while of NSGA-II and SPEA2 are set to be the same as the population size, 100.

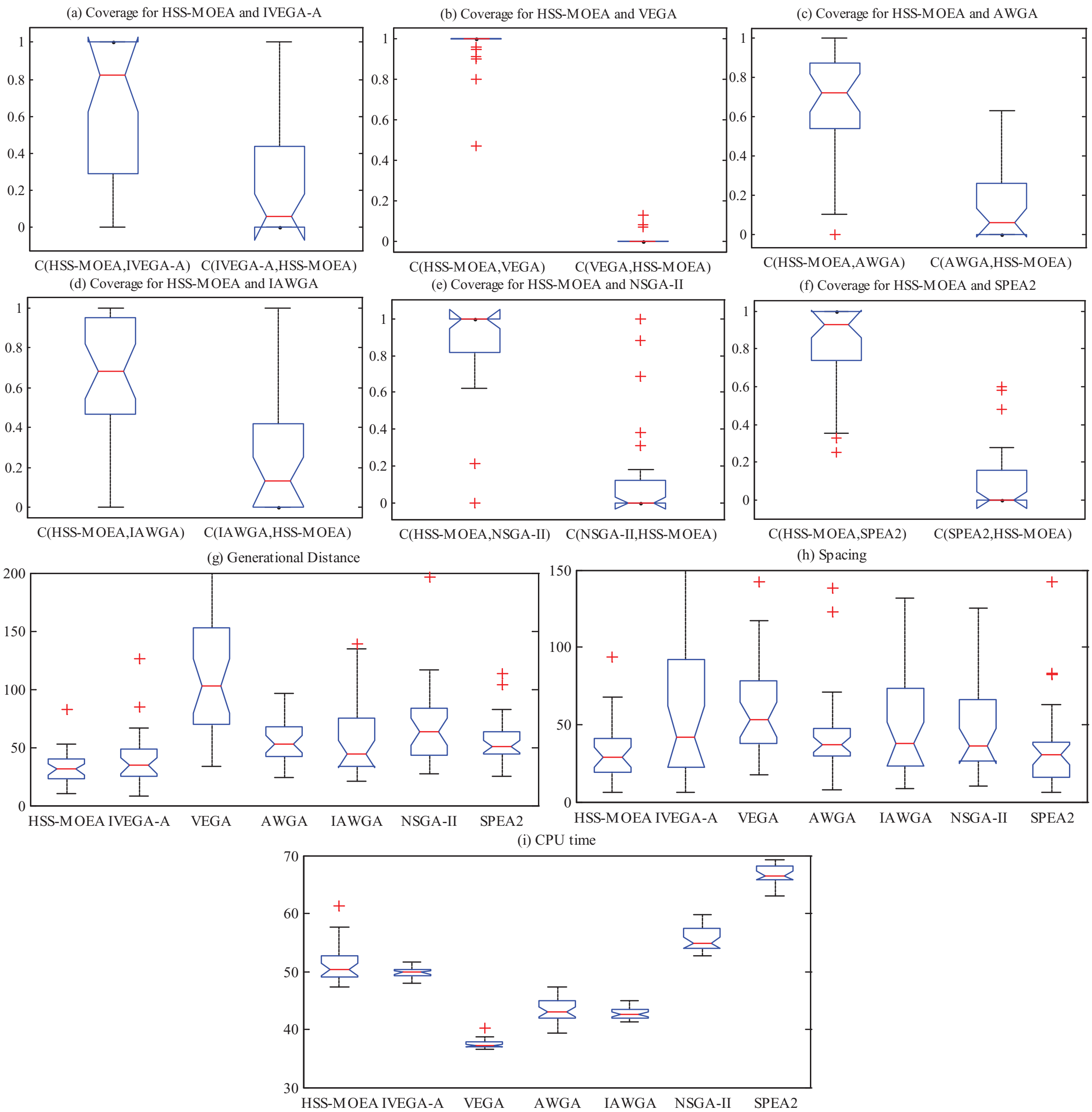

Fig. 2: *C*, *GD*, *SP* and CPU times by 7 methods

Let S_j be a solution set for each method. *PF** is a known reference Pareto solutions. In this study, *PF**in this study comes from combining all of the obtained Pareto set with 30 runs by 7 methods. The following three performance measures are considered.

Coverage $C(S_1,S_2)$ is the percent of the individuals in S_2 which are weakly dominated by S_1 [11]. The larger $C(S_1,S_2)$ means that S_1 outperforms S_2 in convergence.

Generational distance $GD(S_j)$ finds an average minimum distance of the solutions of S_j from *PF** [12]. The smaller *GD* of S_j means better S_j in approaching *PF**.

Spacing $SP(S_j)$ is the standard deviation of the closest distances of individuals by S_j [13]. Smaller *SP* means better distribution performance.

The *C*, *GD* are used to verify convergence performance while *SP* is used to check the distribution performance.

The Fig. 2 shows the numerical comparison of the box-and-whisker plots for *C*, *GD* and *SP* and the CPU times by 7 methods. From Fig. 2a-f, it is easy to see that the HSS-MOEA is better than other 6 methods on *C* measure. The *GD* measure as shown in Fig. 2g also indicates that HSS-MOEA can get slightly smaller than IVEGA-A and smaller value than other 5 methods. Therefore, HSS-MOEA is better than IVEGA-A, VEGA, AWGA, IAWGA as well as famous NSGA-II, and SPEA2.

The distribution performance *SP*(as shown in Fig. 2h) indicates that HSS-MOEA is slightly better than SPEA2, and obviously better than left 5 methods (IVEGA-A, VEGA, AWGA, IAWGA, NSGA-II).

From the comparisons of CPU time as shown in Fig. 2i, it is clear that HSS-MOEA is much faster than NSGA-II and SPEA2 while close to IVEGA-A, and much computational time than VEGA, AWGA and IAWGA.

In general, the convergence and distribution performance of HSS-MOEA is better than IVEGA-A, VEGA, AWGA, IAWGA as well as famous NSGA-II and SPEA2, and the efficiency is obviously better than NSGA-II and SPEA2. Such better convergence and distribution performances should mainly attribute to the hybrid sampling strategy of VEGA's preference for the edge region of the Pareto front and PDDR-FF's tendency converging toward the center area of the Pareto front. They preserve better performances both in efficacy and efficiency. Especially, HSS-MOEA can also keep diversity evenly without special distribution mechanisms like NSGA-II and SPEA2.

5. Conclusions

In this study, a HSS-MOEA approach was proposed to solve MOOP. A new PDDR-FF was proposed to evaluate the individuals. The proposed method tactfully conbined the advantages of VEGA and PDDR-FF based archive strategies. PDDR-FF provides not only a clear classification between nondominated solution and dominated solution, but also a clear difference among solutions locating along the Pareto front. The sampling strategy of VEGA has a preference for the edge region of the Pareto front and the PDDR-FF-based sampling strategy has the tendency converging toward the center area of the Pareto front. Thus hybrid sampling stratigy preserves that the solutions approach to the true Pareto front as close as possible in various directions. Therefore, the proposed method could preserve both the convergence and distribution performances. Numerical comparisons indicated that HSS-MOEA was better than IVEGA-A in efficacy while the efficiency was closely equivalent, and both convergence and distribution performance were also better than NSGA-II and SPEA2 as well as VEGA,AWGA and IAWGA, furethermore, the efficiency was obviously better than NSGA-II and SPEA2.

Acknowledgements

This work is partly supported by the Grant-in-Aid for Scientific Research (C) of Japan Society of Promotion of Science (JSPS: No. 245102190001), National Science Council (NSC 101-2811-E-007-004; NSC100-2628-E-007-017-MY3) and National Tsing Hua University (101N2073E1).

References

1. Gen, M. & Cheng, R. (1997). *Genetic Algorithms and Engineering Design*, John Wiley and Sons, New York.
2. Deb, K. (2001). *Multi-objective Optimization using Evolutionary Algorithms.* John Wiley and Sons, Chichester.
3. Gen, M., Cheng, R. & Lin, L. (2008). *Network Models and Optimization: Multiobjective Genetic Algorithm Approach*, Springer, London.
4. Schaffer, J. (1985). Multiple objective optimization with vector evaluated genetic algorithms, *Proceedings of the 1st International Conference on Genetic Algorithms*. 93–100, Hillsdale, NJ, USA.
5. Gen, M. & Cheng, R. (2000). *Genetic Algorithms and Engineering Optimization*, John Wiley & Sons, New York.
6. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197.
7. Zitzler, E., Laumanns, M., & Thiele, L. (2001). SPEA2: Improving the Strength Pareto Evolutionary Algorithm, Technical Report 103, Computer Engineering and Communication Networks Lab (TIK).
8. Ho, S., Shu, L., & Chen, J. (2004). Intelligent evolutionary algorithms for large parameter optimization problems, *IEEE Transactions on Evolutionary Computation*, 8(6), 522–541.
9. Zhang, W. & Gen, M. (2010). Process planning and scheduling in distributed manufacturing system using multiobjective genetic algorithm, *IEEJ Transactions on Electrical and Electronic Engineering*, 5(1), 62–72.
10. Zhang, W. & Fujimura, S. (2012). Multiobjective process planning and scheduling using improved vector evaluated genetic algorithm with archive, *IEEJ Transactions on Electrical and Electronic Engineering*, 7(3), 258–267.
11. Zitzler, E. & Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation*, 3(4), 257–271.
12. Van Veldhuizen, D. (1999). *Multiobjective Evolutionary Algorithms*: *Classifications, Analyses, and New Innovations*, Ph.D. thesis, Air Force Inst. of Tech., Ohio.
13. Schott, J. (1995). *Fault tolerant design using single and multicriteria genetic algorithm optimization*, Master's thesis, Air Force Inst. of Tech., Ohio.